NIH ERA STATUS REPORT

March 1998

Table of Contents

Introduction

This represents a portion of the fifth status report on NIH electronic research administration (ERA) and reinvention activities that appeared on the NIH Web site in May 1997 and NIH Guide for Grants and Contracts, Volume 26, Number 17, May 23, 1997. The excerpted portion on ERA presented in the current document has been further updated to include development and deployment activities through August 1997. Reports prior to May were published in the NIH Guide for Grants and Contracts on the following dates:

- Volume 23, Number 44, December 16, 1994;
- Volume 24, Number 14, April 14, 1995;
- Volume 24, Number 40, November 24, 1995; and
- Volume 25, Number 23, July 12, 1996.

Electronic Research Administration

Improving stewardship is a prime ongoing objective at the NIH in the administration of grants and contracts awarded to support research at universities and other research facilities around the Nation. Improving the efficiency of various administrative business processes and introducing more efficient ways to communicate relevant information between the NIH and grantee/contractor organizations will maximally leverage appropriations devoted to research, and thereby contribute directly to the productivity of the nation's medical research enterprise.
In addition to pursuing reinvention activities to improve research administration processes, the phenomenal advances and almost constant changes in Information Technology (IT) are also being exploited at the NIH to improve stewardship of awards.

NIH is currently devoting substantial IT-related resources in the design, development and deployment of an electronic research administration (ERA) system. This ERA system will greatly facilitate preparation of grant applications by research investigators, processing of applications by NIH staff, and management of awards by grantee organizations and NIH staff.

The ERA system will eventually place the entire grants administration life cycle of business processes within a client-server common file database. Using this database technology will permit the NIH to maintain timely, fully electronic communication with extramural grantee "business partners." The system will be made fully secure using state-of-the-art encryption methodology. Access will be limited to authorized applicants, awardees and NIH staff, who could each review and add information as required. Proposed components of the system include submission and review of competitive grant applications; maintenance of contact information for all NIH grantee organizations and investigators; a status system by which authorized users can determine the status of active and pending grants (including review dates, scores, and application critiques); electronic notification to the grantee organization of grant award; and a full function search interface to review awarded research project abstracts. Other post-award reporting features will include financial, research progress, research trainee appointments, and inventions.

The NIH Commons

To accommodate these proposed components, the NIH has begun a formal software development process to enhance both the discrete exchange of essential information between NIH and applicant organizations. Central to these improvements is the concept of the "Commons," the information interface where NIH and the grantee community will conduct their business electronically. The systematic development and implementation of the Commons will be crucial to the success of NIH's ERA efforts.

The Commons is being designed around system components that interact with an Oracle™ database. Interaction with the database will provide functionality to support extramural grants administration business processes. To accommodate the broadest set of extramural user requirements, the Commons will interface with the extramural community via secure interactive Internet sites. Software behind these sites provide all necessary functionality for services such as security, auditing, record submission, updating, modification, and in some instances deletion. The rapid advances in Internet technology have made this approach highly feasible, while cost-effective and well received by the broadest spectrum of extramural grantee/contractor organizations.

NIH is taking a well-documented, methodical approach to designing the portions of the ERA system that will support these varied business processes. The components of the ERA system that have been built and are either undergoing pilot testing, or have been fully deployed as summarized below. For the majority of the ERA functions described below, CY 97 is a pilot year (pilot testing with ten institutions, as described below); CY 98 is a transition year (expanded pilot testing with Federal Demonstration Partnership institutions); and CY 99 is the goal for full production implementation.

Commons Data Standards and Formats

There exist two principle objectives of the Commons design that will ensure that the system is compatible with the greatest number of grantee organizations. First, all Commons software applications are being built upon well-documented data standards. The most notable of these is the public standard for grants administration, the American National Standards Institute (ANSI) 194 transaction set. By invoking the use of this data dictionary for all grants administration transmissions, both the grantee organization, sending the information, and the government agency, receiving the information, will be assured of successful communication. Similar standards are being employed for all business process transactions being developed in the Commons system.

The second Commons design objective is to maintain flexibility in the interface requirements available to Commons grantee organization users. In order for the NIH to accommodate a range of grantee organizations with disparate information technology infrastructures, the Commons will support three well-known transmission formats. Grantee organizations with substantial database systems wishing to interface with the Commons via computer-to-computer transmissions will be

able to employ two standard datastream formats: Electronic Data Interchange (EDI), and/or Hypertext Markup Language (HTML). By contrast, other organizations that do not wish to employ computer-to-computer transmissions will be able to exchange information using basic Internet connectivity, using standard Web browser technology exchanging browser-derived HTML files. Details as to how these formats have been incorporated into specific Commons interfaces are described below.

Electronic Submission of Competing Grant Applications

Electronic Data Interchange (EDI)

Under a Department of Energy (DOE) Cooperative Agreement, the NIH along with several Department of Defense (DoD) agencies participated over the last several years in a pilot study to test a new system for the submission of grant application information. These agencies and ten research institutions are continuing to test the EDI standards developed collaboratively by the Federal agencies. Key administrative information in R01 grant applications, such as face page information, scientific abstract, certain budget items, and personal data for the Principal Investigator, will be submitted directly into NIH's database, without intervening paper copies or manual re-keying of data. This pilot implementation commenced in October 1996, will continue through 1997 and will be expanded in 1998 to include more aspects of the applications and more grantee organizations. The initial DOE pilot permitted the submission of only the application shell information (PHS form 398 pages AA, BB, DD, EE, Abstract, and Specific Aims). The FDP pilot will expand this information to include other pages (GG, HH, KK) as well as by allowing the Research Plan aspect of the application to be submitted as an Adobe PDF™ file. Other information required for the submission of the competitive application will be provided through linking to another portion of the Commons/IMPAC II databases that provide for Professional and Institutional Profile information (see Admin. System for details)

HyperText Markup Language (HTML)

The first pilot implementation of the competitive application involves transmission of data via EDI formatted files. In an effort to provide grantee organizations with alternative modes of transmission of the same data, the NIH is currently piloting software that will allow for the receipt of competitive application information via an HTML-formatted data stream. HTML is the language used to format data for use and visualization on the World Wide Web. Thus, allowing for receipt of grant application information in this format affords grantee organizations who are developing research administration systems centered on Internet and HTML to retain data in this common format as they prepare to submit information to the NIH. It is important to emphasize that both EDI and HTML- formatting datastreams will be based strictly on the same standard set of data elements. NIH began demonstration of the HTML prototype software in March 1997. By the Spring of 1998, FDP organizations will be able to submit the same extent of information via

an HTML datastream as by EDI datastream, including the Research Plan formatted as an Adobe PDF™ file.

Electronic Streamlined Noncompeting Award Process (e-SNAP)

e-SNAP Submission Via Interactive Web

FY 95, NIH instituted a simplified noncompeting award process (SNAP) for the majority of noncompeting continuation awards. Under SNAP, which applies to awards under the Expanded Authorities and Federal Demonstration Partnership, certain components of the noncompeting application are not required if there are no significant changes. (For more information on SNAP, see the NIH Guide for Grants and Contracts, October 27, 1995 and July 5, 1996.)

An ERA version of the SNAP process has been pilot tested to the DOE cooperative agreement organizations since November 1996. "e-SNAP" is an interactive World Wide Web- based site for electronic submission of the SNAP information. Using the interface, authorized grantees are able to prepare for submission all required information. Then, through a separate site, grantee organization administrative official(s) can approve and then finally submit the application to the NIH to initiate the noncompeting award process. Upon receipt of the e-SNAP submission, NIH staff evaluate the electronic application and if approved generate an electronic Notice of Grant Award back to the grantee and grantee organization administrative official(s). The next phase of e-SNAP deployment will occur in Spring, 1998 to the FDP participating organizations. This version of the interactive Web-based version of e-SNAP will include expanded functionality as suggested by the DOE cooperative agreement pilot organizations, including an enhanced functionality for routing of applications for approval by the grantee organization prior to submission to the NIH.

A final major phase of functionality enhancement for non-competing award applications will be to include an interactive Web-based version of the software that will permit submission of even applications that do not qualify for Streamlined Noncompeting Award Applications. These enhancements will allow for the submission of, for example, complex Centers or Program Project grants will not be available for testing until 1999.

e-SNAP submission via EDI or HTML Datastream

As an adjunct to e-SNAP via an interactive Web interface, as part of the FDP pilot deployment in the Spring of 1998, grantee organizations will be given the alternative to submit the SNAP information using either EDI or HTML-formatted datastreams.

Electronic Reporting of Trainee Activities

Organizations that receive National Research Service Act (NRSA) Institutional research training grants must report on the appointment of trainee(s) supported under the grant. An ERA system interface is now being piloted that provides for the electronic submission of trainee appointment information. Like e-SNAP, the trainee appointment system takes advantage of the user-friendly benefits of the World Wide Web. In this case, the grant Training Program Director will submit required information through the secure Web interface. The pilot deployment of this portion of the ERA system started at the beginning of FY 96 with test data. "Live" data was received from eight grantee organizations in mid-FY 96, with NIH staff issuing electronic approvals. Expansion to include the FDP organizations will occur in Spring, 1998. As of now, there are no plans to receive this information via either EDI or HTML datastreams.

Application and Award Status System

One of the obvious benefits of electronic communication is the ability to exchange time-sensitive information in a timely manner. Once critical administrative decisions or updates in information occur, the results of these actions can be communicated rapidly. This type of functionality for NIH extramural grants administration will be provided through the Commons Application and Award Status interface. In the first phase of the pilot test of Status, which began in March 1997, authorized users from the DOE cooperative agreement organizations were able to log onto a secure Web site where they could review basic aspects of the status of applications sent to the NIH, including pending review, review outcome, pending Advisory Council action, and award status. Grantee organization officials are be able to view pending actions for all applications originating from their organization, while individual investigators are able to view only their applications.

The second phase of deployment of this portion of the Commons, to follow in April 1998, will provide an expanded number of status milestones. In addition, the interface will enable research investigators to receive the text of Summary Statement(s) as well as priority score(s) and percentile rankings for any of their reviewed but yet to be awarded applications. For this information, it is important to note that Summary Statements, priority scores, and percentile ranking will be available only to the Principal Investigator: Administrative Officials will not have access to this information.

Commons Admin. System

<u>ERA Accounts Administration</u>

On an annual basis, the NIH engages approximately 2,000 individual grantee organizations in grants/contract administration business processes. Some of these organizations employ 100's to 1,000's of investigators. Accordingly, for the NIH to be able to establish secure authenticated electronic interactions and then track use and administration of accounts for this potentially

overwhelming population of end users, the Commons design has employed a hierarchical administration of ERA user accounts. The Accounts Administration will be administered through the Commons Admin. System interface. It will permit senior administrative officials (Signing Officials) at each grantee organization to establish an account for the organization proper. Once this account is in place the Signing Officials will be able to create secondary accounts for administrative officials, who can in turn create additional accounts for administrative staff and scientific staff. This will permit the administrative officials - the individuals who have the best working knowledge of the workforce at their organization - to create and terminate accounts consistent with the status of investigators and administrative staff. Moreover, this hierarchy will establish the means by which applications will be approved prior to submission to the NIH.

Professional and Institutional Profile Administration

In addition to accounts administration, the Commons will support the establishment, monitoring and updating of two types of information profiles in order to relieve Administrative Officials, Principal Investigators and grant key personnel from having to re-key information. Principal Investigators will create and then be able to access a Professional Profile File (PPF) to update their contact information and biosketch information. In a similar way, authorized Grantee Organization officials will be able to update the organizational contact information in an Institutional Profile File (IPF). These actions will preclude re-keying of such information for each submitted application, as must now be done with paper submissions.

Electronic Invention Reporting

The requirement for reporting of information pursuant to inventions derived from Federal funds is mandated in the Bayh-Dole Act. To support this requirement the NIH has developed an ERA system dubbed "Edison", designed to receive, store, sort, and provide reports on invention, patent, licensing and invention utilization. As the first secure interactive Web site developed in ERA, it has been deployed in a full production version since the beginning of FY 96. The mission of ERA has been realized and can be documented by the Edison system. After one year, almost 50 percent of the grantee organizations that report inventions routinely to the NIH are using Edison. For each invention this represents the reduction from typically 15 cycles of paper correspondence to 3, dramatically shortening reporting time and effort, as well as making more information available in a usable format for grants administrators. In addition to continued recruitment of grantee organizations to use Edison, the next phase for this ERA system component will be to include other Federal research agencies in the deployment. Six additional agencies, including NSF, FDA, CDCP, USDA, NOAA, and have signed formal Memoranda of Understanding through which they have committed to use the Edison system to support invention reporting by their grantees. The deployment of "Interagency Edison" commenced in the Summer of 1997.

CRISP on the Web

The NIH awards database, Computer Retrieval of Information on Scientific Projects (CRISP), is a collection of research project abstracts and grant related information that spans nearly 25 years. It is accessible via a "gopher" interface, which is limited to relatively simple queries. As part of the ERA initiative, a Web-based CRISP interface will allow full text searching of research project abstracts and grantee organization information. This will improve the ability to retrieve the most relevant sources of information to facilitate further research through inquires by researchers, to provide useful orientation for the public, and to improve the efficiency of responses to inquires issued by other agencies and Congress. The new CRISP Web interface will be pilot-tested with NIH staff commencing in the Spring of 1998. Given the popularity of the interface and the fact that the information can be freely disseminated to the public, efforts will be made to fully deploy the CRISP interface to the public by Summer 1998.

Commons deployment

DOE Cooperative Agreement Pilot Deployment

The Commons was first deployed on a development computer platform in November 1996. This platform was made accessible to 10 grantee organizations who had been selected to assist in the first pilot deployment. The organizations included:

- Fred Hutchinson Cancer Center
- Univ. of California, Los Angeles
- Pennsylvania State University
- University of Notre Dame
- Florida A&M University
- Ohio State University
- North Carolina State University
- Baylor College of Medicine
- Duke University
- Massachusetts Institute of Technology

This selection was based on participation in a DOE Cooperative Agreement, in which NIH was formally involved. As Commons functionality was completed and introduced on the Commons, these 10 grantee organizations were invited to access and test the functionality. Deployment of the NIH Commons in this way has gone well, with grantee organizations providing significant useful feedback to enhance Commons usability.

FDP Pilot Deployment

The next phase of the deployment is currently scheduled to commence at the end of the first quarter of CY1998. At that time use of the Commons will be expanded to include approximately 60 grantee organizations participating in the Federal Demonstration Partnership (FDP). As with the deployment to DOE cooperative agreement organizations, the FDP participating organizations will be invited to establish accounts to use the Commons for the purposes of submission of information, accessing the status of applications previous submitted to NIH, and establishment of both professional profiles for grantees, as well as a profile for the related institution. Initially, transmission will be comprised of "test" information, followed by "live" data exchanges. The

distinction between these types of transmission will be the ability for the NIH Commons to successful upload the transmitted information to the NIH enterprise database system, IMPAC II. During the test transmissions the information received by the Commons will be examined and validated. Uploading to IMPAC II will follow, with further validation prior to having the record become part of the enterprise database. Once this uploading has been fully tested, receipt of "live" data will proceed. With initial deployment to FDP organizations in April, 1998 for receipt of "test" information, it is hoped that "live" submissions will occur in the June/July timeframe.

Throughout the FDP deployment, institutions will be submitting paper applications in parallel with electronic applications to ensure no interruption in application processing, even should any problems arise in the processing of the electronic applications.

Pre-Production Rollout

By the end of CY1998, based on the FDP pilot deployment, the NIH will have considerable feedback and experience necessary to further refine the Commons interfaces. With this input as as a basis for software improvement, in the first quarter of 1999 the Commons deployment will be scaled up as a Pre-Production Rollout. What is anticipated is that the applications and status functionality will be expanded such that approximately 200 grantee organizations will be invited to participate in further deployment. The exact implementation plan of this Pre-Production Rollout has not yet been finalized. To a large extent, the timing of this phase of deployment will be determined by the successes as well as any problems that arise commensurate with FDP deployment.

Final Production Rollout

The current deployment plan calls for a final Production Rollout commencing in the early part of 2000. As with the Pre-Production Rollout, as of now this milestone should be considered approximate. The exact timeframe will come into focus depending on the status and success of precedent pilots and rollout.

Critical Short-term Milestones

<u>Deployment of the Commons Production Platform</u>

In order for the deployment of the NIH Commons consistent with the plan summarized above, two critical short-term milestones have been, or are in the final stages of being met. First, the Production Commons platform has been activated. The production platform is a Digital Equipment Corporation (DEC) Alpha™ computer located at, and supported fully by the NIH Division of Computer Resources and Technology (DCRT). The computer is in place and the first

instance of the Commons database and interfaces have been placed in this production environment. With the Commons interface migrated to the DEC with some functional interfaces in place, once development and testing of additional functionality is complete these interfaces will follow, consistent with the pilot deployment schedule described above.

Replication of the Commons with IMPAC II

Throughout the DOE Cooperative Agreement Pilot deployment, the Commons database and IMPAC II databases have not been connected. That is, as pilot information is submitted by the 10 grantee organizations, that information has not been transferred to IMPAC II. In order for this transfer to occur, a set of complex replication functions have to take place, whereby the Commons-derived information synchronizes properly with records residing in IMPAC II. For this complex activity to occur properly close coordination of software coding and timing mechanisms have had to be defined for both Commons and IMPAC II systems.

The NIH contractors responsible for design, development, testing, and deployment of the Commons and IMPAC II, TYC Associates and R.O.W. Sciences, respectively have worked very closely toward meeting this critical requirement. The replication process is now in final testing, with transfer of information between the systems expected to be fully implemented commensurate with deployment of the Commons to the FDP organizations in April 1998.